1600

RAW SEQUENCE LISTING
PATENT APPLICATION: **US/09/973,322A**

DATE: 04/30/2003
TIME: 11:56:35

Input Set : **A:\DYOU17.001CP1SEQLIST.TXT**
Output Set: **N:\CRF4\04302003\I973322A.raw**

```
 4 <110> APPLICANT: Hope, Ralph Graham
 5         McLauchlan, John
 7 <120> TITLE OF INVENTION: VIRAL THERAPEUTICS
10 <130> FILE REFERENCE: DYOU17.001CP1
12 <140> CURRENT APPLICATION NUMBER: US 09/973,322A
13 <141> CURRENT FILING DATE: 2001-10-09
15 <150> PRIOR APPLICATION NUMBER: US 09/201,916
16 <151> PRIOR FILING DATE: 1998-12-01
18 <150> PRIOR APPLICATION NUMBER: GB 9825951.8
19 <151> PRIOR FILING DATE: 1998-11-26
21 <160> NUMBER OF SEQ ID NOS: 22
23 <170> SOFTWARE: FastSEQ for Windows Version 4.0
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 630
27 <212> TYPE: DNA
28 <213> ORGANISM: Hepatitis C Virus
30 <220> FEATURE:
31 <221> NAME/KEY: CDS
32 <222> LOCATION: (43)...(630)
34 <400> SEQUENCE: 1
35 ggtgcttgcg agtgccccgg gaggtctcgt agaccgtgca cc atg agc acg aat    54
36                                                  Met Ser Thr Asn
37                                                   1
39 cct aaa cct caa aga aaa acc aaa cgt aac acc aac cgt cgc cca cag   102
40 Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln
41  5              10                  15                  20
43 gac gtt aag ttc ccg ggt ggc ggt cag atc gtt ggt gga gtt tac ttg   150
44 Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr Leu
45              25                  30                  35
47 ttg ccg cgc agg ggc cct aga ttg ggt gtg cgc gcg acg agg aag act   198
48 Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala Thr Arg Lys Thr
49          40                  45                  50
51 tcc gag cgg tcg caa cct cga ggt aga cgt cag cct atc ccc aag gca   246
52 Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro Ile Pro Lys Ala
53          55                  60                  65
55 cgt cgg ccc aag ggc agg aac tgg gct cag ccc ggg tat cct tgg ccc   294
56 Arg Arg Pro Lys Gly Arg Asn Trp Ala Gln Pro Gly Tyr Pro Trp Pro
57      70                  75                  80
59 ctc tat ggc aat gag ggt tgc ggg tgg gcg gga tgg ctc ctg tcc ccc   342
60 Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp Leu Leu Ser Pro
61  85                  90                  95                  100
63 agt ggc tct cgg cct agt tgg ggc ccc aac gac ccc cga cgt agg tcg   390
64 Ser Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro Arg Arg Arg Ser
```

RAW SEQUENCE LISTING   DATE: 04/30/2003
PATENT APPLICATION: **US/09/973,322A**   TIME: 11:56:35

Input Set : **A:\DYOU17.001CP1SEQLIST.TXT**
Output Set: **N:\CRF4\04302003\I973322A.raw**

```
65                    105                  110                    115
67 cgc aat ttg ggt aag gtc atc gat acc ctt acg tgc ggc ttc gtc gat    438
68 Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Val Asp
69            120                  125                  130
71 ctc atg ggg tac ata ccg ctc gtc ggc gcc cct ctt aga ggc gct gcc    486
72 Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Arg Gly Ala Ala
73            135                  140                  145
75 agg gcc ctg gcg cat ggc gtc cgg gtt ctg gaa gac ggt gtg aac tat    534
76 Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr
77        150                  155                  160
79 gca aca ggt aac ctt cct ggt tgc tct ttc tct atc ttc ctt ctg gcc    582
80 Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala
81 165                  170                  175                  180
83 ctg ctc tct tgc ctg act gtg ccc gct tca gcc tac caa gtg cgc aac    630
84 Leu Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Gln Val Arg Asn
85                  185                  190                  195
89 <210> SEQ ID NO: 2
90 <211> LENGTH: 60
91 <212> TYPE: DNA
92 <213> ORGANISM: Hepatitis C Virus
94 <220> FEATURE:
95 <221> NAME/KEY: CDS
96 <222> LOCATION: (1)...(60)
97 <223> OTHER INFORMATION: Corresponds to aa 125 to 144 of SEQ ID. No. 1
99 <400> SEQUENCE: 2
100 acc ctt acg tgc ggc ttc gtc gat ctc atg ggg tac ata ccg ctc gtc    48
101 Thr Leu Thr Cys Gly Phe Val Asp Leu Met Gly Tyr Ile Pro Leu Val
102 1             5                  10                  15
104 ggc gcc cct ctt                                                    60
105 Gly Ala Pro Leu
106             20
109 <210> SEQ ID NO: 3
110 <211> LENGTH: 18
111 <212> TYPE: DNA
112 <213> ORGANISM: Hepatitis C Virus
114 <220> FEATURE:
115 <221> NAME/KEY: CDS
116 <222> LOCATION: (1)...(18)
117 <223> OTHER INFORMATION: Corresponds to aa 161-166 of SEQ ID. No. 1
119 <400> SEQUENCE: 3
120 ggt gtg aac tat gca aca                                            18
121 Gly Val Asn Tyr Ala Thr
122 1             5
125 <210> SEQ ID NO: 4
126 <211> LENGTH: 1900
127 <212> TYPE: DNA
128 <213> ORGANISM: Human
130 <220> FEATURE:
131 <221> NAME/KEY: misc_feature
```

RAW SEQUENCE LISTING   DATE: 04/30/2003
PATENT APPLICATION: **US/09/973,322A**   TIME: 11:56:35

Input Set : **A:\DYOU17.001CP1SEQLIST.TXT**
Output Set: **N:\CRF4\04302003\I973322A.raw**

```
    132 <222> LOCATION: (1)...(1900)
    133 <223> OTHER INFORMATION: n = A,T,C or G
    135 <400> SEQUENCE: 4
    136 cgtcttcggg acgcgcccgc tcttcgcctt tcgctgcagt ccgtcgattt ctttctccag 60
    137 gaagaaaaat ggcatccgtt gcagttgatc cacaaccgag tgtggtgact cgggtggtca 120
    138 acctgccctt ggtgagctcc acgtatgacc tcatgtcctc agcctatctc agtacaaagg 180
    139 accagtatcc ctacctgaag tctgtgtgtg agatgscaga gaacggtgtg aagaccatca 240
    140 cctccgtggc catgaccagt gctctgccca tcatccagaa gctagagccg caaattgcag 300
    141 ttgccgatac ctatgcctgt aaggggctag acaggattga ggagagactg cctattctga 360
    142 atcagccatc aactcagatt gttgccaatg ccaaaggcgc tgtgactggg gcaaaagatg 420
W-->143 ctgtgacgac tactgtgact ggggccaagg attctgtngc cagcacgatc acaggggtga 480
    144 tggacaagac caaaggggca gtgactggca gtgtggagaa gaccaagtct gtggtcagtg 540
    145 gcagcattaa cacagtcttg gggagtcgga tgatgcagct cgtgagcagt ggcgtagaaa 600
    146 atgcactcac caaatcagag ctgttggtag aacagtacct ccctctcact gaggaagaac 660
    147 tagaaaaaga agcaaaaaaa gttgaaggat ttgatctggt tcagaagcca agttattatg 720
    148 ttagactggg atccctgtct accaagcttc actcccgtgc ctaccagcag gctctcagca 780
    149 gggttaaaga agctaagcaa aaaagccaac agaccatttc tcagctccat tctactgttc 840
    150 acctgattga atttgccagg aagaatgtgt atagtgccaa tcagaaaatt caggatgctc 900
    151 aggataagct ctacctctca tgggtagagt ggaaaaggag cattggatat gatgatactg 960
    152 atgagtccca ctgtgctgag cacattgagt cacgtactct tgcaattgcc cgcaacctga 1020
    153 ctcagcagct ccagaccacg tgccacaccc tcctgtccaa catccaaggt gtaccacaga 1080
    154 acatccaaga tcaagccaag cacatggggg tgatggcagg cgacatctac tcagtgttcc 1140
    155 gcaatgctgc ctcctttaaa gaagtgtctg acagcctcct cacttctagc aaggggcagc 1200
    156 tgcagaaaat gaaggaatct ttagatgacg tgatggatta tcttgttaac aacacgcccc 1260
    157 tcaactggct ggtaggtccc ttttatcctc agctgactga gtctcagaat gctcaggacc 1320
    158 aaggtgcaga gatggacaag agcagccagg agacccagcg atctgagcat aaaactcatt 1380
    159 aaacctgccc ctatcactag tgcatgctgt ggccagacag atgacacctt ttgttatgtt 1440
    160 gaaattaact tgctaggcaa ccctaaattg ggaagcaagt agctagtata aaggccctca 1500
    161 attgtagttg tttccagctg aattaagagc tttaaagttt ctggcattag cagatgattt 1560
    162 ctgttcacct ggtaagaaaa gaatgatagg cttgtcagag cctatagcca gaactcagaa 1620
    163 aaaattcaaa tgcacttatg ttctcattct atggccattg tgttgcctct gttactgttt 1680
    164 gtattgaata aaaacatctt catgtgggct ggggtagaaa ctggtgtctg ctctggtgtg 1740
    165 atctgaaaag gcgtcttcac tgctttatct catgatgctt gcttgtaaaa cttgatttta 1800
    166 gttttcatt tctcaaatag gaatactacc tttgaattca ataaaattca ctgcaggata 1860
    167 gaccagttna gnagcaaaca nncangtaca cnnaaganac                        1900
    169 <210> SEQ ID NO: 5
    170 <211> LENGTH: 437
    171 <212> TYPE: PRT
    172 <213> ORGANISM: Human
    174 <220> FEATURE:
    175 <221> NAME/KEY: VARIANT
    176 <222> LOCATION: (1)...(437)
    177 <223> OTHER INFORMATION: Xaa = Any Amino Acid
    179 <400> SEQUENCE: 5
    180 Met Ala Ser Val Ala Val Asp Pro Gln Pro Ser Val Val Thr Arg Val
    181   1               5                  10                  15
    182 Val Asn Leu Pro Leu Val Ser Ser Thr Tyr Asp Leu Met Ser Ser Ala
    183              20                  25                  30
    184 Tyr Leu Ser Thr Lys Asp Gln Tyr Pro Tyr Leu Lys Ser Val Cys Glu
```

RAW SEQUENCE LISTING   DATE: 04/30/2003
PATENT APPLICATION: **US/09/973,322A**   TIME: 11:56:35

Input Set : **A:\DYOU17.001CP1SEQLIST.TXT**
Output Set: **N:\CRF4\04302003\I973322A.raw**

```
         185            35                        40                        45
W-->   186 Met Xaa Glu Asn Gly Val Lys Thr Ile Thr Ser Val Ala Met Thr Ser
       187         50                    55                    60
       188 Ala Leu Pro Ile Ile Gln Lys Leu Glu Pro Gln Ile Ala Val Ala Asp
       189 65                    70                    75                    80
       190 Thr Tyr Ala Cys Lys Gly Leu Asp Arg Ile Glu Glu Arg Leu Pro Ile
       191                 85                    90                    95
       192 Leu Asn Gln Pro Ser Thr Gln Ile Val Ala Asn Ala Lys Gly Ala Val
       193                100                   105                   110
       194 Thr Gly Ala Lys Asp Ala Val Thr Thr Thr Val Thr Gly Ala Lys Asp
       195         115                   120                   125
       196 Ser Val Ala Ser Thr Ile Thr Gly Val Met Asp Lys Thr Lys Gly Ala
       197         130                   135                   140
       198 Val Thr Gly Ser Val Glu Lys Thr Lys Ser Val Val Ser Gly Ser Ile
       199 145                   150                   155                   160
       200 Asn Thr Val Leu Gly Ser Arg Met Met Gln Leu Val Ser Ser Gly Val
       201                165                   170                   175
       202 Glu Asn Ala Leu Thr Lys Ser Glu Leu Leu Val Glu Gln Tyr Leu Pro
       203         180                   185                   190
       204 Leu Thr Glu Glu Glu Leu Glu Lys Glu Ala Lys Lys Val Glu Gly Phe
       205         195                   200                   205
       206 Asp Leu Val Gln Lys Pro Ser Tyr Tyr Val Arg Leu Gly Ser Leu Ser
       207         210                   215                   220
       208 Thr Lys Leu His Ser Arg Ala Tyr Gln Gln Ala Leu Ser Arg Val Lys
       209 225                   230                   235                   240
       210 Glu Ala Lys Gln Lys Ser Gln Gln Thr Ile Ser Gln Leu His Ser Thr
       211                245                   250                   255
       212 Val His Leu Ile Glu Phe Ala Arg Lys Asn Val Tyr Ser Ala Asn Gln
       213         260                   265                   270
       214 Lys Ile Gln Asp Ala Gln Asp Lys Leu Tyr Leu Ser Trp Val Glu Trp
       215         275                   280                   285
       216 Lys Arg Ser Ile Gly Tyr Asp Asp Thr Asp Glu Ser His Cys Ala Glu
       217         290                   295                   300
       218 His Ile Glu Ser Arg Thr Leu Ala Ile Ala Arg Asn Leu Thr Gln Gln
       219 305                   310                   315                   320
       220 Leu Gln Thr Thr Cys His Thr Leu Leu Ser Asn Ile Gln Gly Val Pro
       221                325                   330                   335
       222 Gln Asn Ile Gln Asp Gln Ala Lys His Met Gly Val Met Ala Gly Asp
       223         340                   345                   350
       224 Ile Tyr Ser Val Phe Arg Asn Ala Ala Ser Phe Lys Glu Val Ser Asp
       225         355                   360                   365
       226 Ser Leu Leu Thr Ser Ser Lys Gly Gln Leu Gln Lys Met Lys Glu Ser
       227         370                   375                   380
       228 Leu Asp Asp Val Met Asp Tyr Leu Val Asn Asn Thr Pro Leu Asn Trp
       229 385                   390                   395                   400
       230 Leu Val Gly Pro Phe Tyr Pro Gln Leu Thr Glu Ser Gln Asn Ala Gln
       231                405                   410                   415
       232 Asp Gln Gly Ala Glu Met Asp Lys Ser Ser Gln Glu Thr Gln Arg Ser
       233         420                   425                   430
```

RAW SEQUENCE LISTING   DATE: 04/30/2003
PATENT APPLICATION: US/09/973,322A   TIME: 11:56:35

Input Set : A:\DYOU17.001CP1SEQLIST.TXT
Output Set: N:\CRF4\04302003\I973322A.raw

```
        234 Glu His Lys Thr His
        235          435
        238 <210> SEQ ID NO: 6
        239 <211> LENGTH: 31
        240 <212> TYPE: PRT
        241 <213> ORGANISM: Artificial Sequence
        243 <220> FEATURE:
        244 <223> OTHER INFORMATION: branched peptide containing residues 5-27 of HCV
        245       core protein
W-->    247 <221> NAME/KEY: VARIANT
        248 <222> LOCATION: (1)...(31)
        249 <223> OTHER INFORMATION: Xaa = Ala or Pro at position 1, and Ile or Asn at
        250       position 12
W-->    252 <400> 6
W-->    253 Xaa Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Xaa Arg Arg Pro Gln
        254  1               5                  10          Xaa          15
        255 Asp Val Lys Phe Pro Gly Gly Lys Lys Lys Lys Lys Lys Lys Ala
        256          20              25              30
        259 <210> SEQ ID NO: 7
        260 <211> LENGTH: 11
        261 <212> TYPE: DNA
        262 <213> ORGANISM: Artificial Sequence
        264 <220> FEATURE:
        265 <223> OTHER INFORMATION: oligonucleotides used to construct HCV core
        266       protein deletion plasmids
        268 <400> SEQUENCE: 7
        269 gctgagatct a                                                11
        271 <210> SEQ ID NO: 8
        272 <211> LENGTH: 29
        273 <212> TYPE: DNA
        274 <213> ORGANISM: Artificial Sequence
        276 <220> FEATURE:
        277 <223> OTHER INFORMATION: oligonucleotides used to construct HCV core
        278       protein deletion plasmids
        280 <400> SEQUENCE: 8
        281 gtaaccttcc tggttgctct tgagatcta                             29
        283 <210> SEQ ID NO: 9
        284 <211> LENGTH: 17
        285 <212> TYPE: DNA
        286 <213> ORGANISM: Artificial Sequence
        288 <220> FEATURE:
        289 <223> OTHER INFORMATION: oligonucleotides used to construct HCV core
        290       protein deletion plasmids
        292 <400> SEQUENCE: 9
        293 gtaacctttg agatcta                                          17
        295 <210> SEQ ID NO: 10
        296 <211> LENGTH: 18
        297 <212> TYPE: DNA
        298 <213> ORGANISM: Artificial Sequence
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 04/30/2003
PATENT APPLICATION: US/09/973,322A   TIME: 11:56:36

Input Set : A:\DYOU17.001CP1SEQLIST.TXT
Output Set: N:\CRF4\04302003\I973322A.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:4; N Pos. 458,1869,1872,1881,1882,1885,1892,1893,1898
Seq#:5; Xaa Pos. 50
Seq#:6; Xaa Pos. 1,12